

The Patent Office Concept House Cardiff Road Newport South Wales NP10 800

I, the undersigned, being an officer duly authorised in accordance with Section 74(1) and (4) of the Deregulation & Contracting Out Act 1994, to sign and issue certificates on behalf of the Comptroller-General, hereby certify that annexed hereto is a true copy of the documents as originally filed in connection with the patent application identified therein.

I also certify that the attached copy of the request for grant of a Patent (Form 1/77) bears an amendment, effected by this office, following a request by the applicant and agreed to by the Comptroller-General.

In accordance with the Patents (Companies Re-registration) Rules 1982, if a company named in this certificate and any accompanying documents has re-registered under the Companies Act 1980 with the same name as that with which it was registered immediately before reregistration save for the substitution as, or inclusion as, the last part of the name of the words "public limited company" or their equivalents in Welsh, references to the name of the company in this certificate and any accompanying documents shall be treated as references to the name with which it is so re-registered.

In accordance with the rules, the words "public limited company" may be replaced by p.l.c., plc, P.L.C. or PLC.

Re-registration under the Companies Act does not constitute a new legal entity but merely subject the company to certain additional company law rules.



Dated 6 April 2004

An Executive Agency of the Department of Trade and Industry

Patents Form 1/77

tents Act 1977 Rule 16)

19HAY03 E808157-1 D03312 P01/7700 0.00-0311351.1 ·

The Patent Office

Cardiff Road

Newport Gwent NP9 1RH

Request for grant of a patent

(See the notes on the back of this form. You can also get an explanatory leaflet from the Patent Office to help you fill in this form)

1. Your reference

GBP290013

- 2. Patent application number (The Patent Office will fill in this part)
- 3. Full name, address and postcode of the or of each applicant (underline all sumames)

2003

0311351.1

Discerna Limited. Babraham Hall Babraham Cambridge Cambridgeshire CB2 4AT United Kingdom

85922 | 400 |

United Kingdom

Patents ADP number (if you know it)

If the applicant is a corporate body, give the country/state of its incorporation

Title of the invention

number

Activity-directed in vitro protein cloning and identification

Marks & Clerk

18001

Wellington House

Name of your agent (if you have one) "Address for service" in the United Kingdom to which all correspondence should be sent (including the postcode)

Patents ADP number (if you know it)

6. If you are declaring priority from one or

more earlier patent applications, give the

country and the date of filing of the or of each of these earlier applications and (if you know it) the or each application

Country

66/68 Hills Road Capibridge CB1 1BH

7271125003

Priority application No (if you know it)

Date of filing (day / month / year)

7. If this application is divided or otherwise derived from an earlier UK application, give the number and the filing date of the earlier application

Number of earlier application

Date of filing (day / month / year)

8. Is a statement of inventorship and of right to grant of a patent required in support of this request? (Answer 'Yes' if:

a) any applicant named in part 3 is not an inventor, or b) there is an inventor who is not named as an applicant, or c) any named applicant is a corporate body. See note (d))

Yes

#### Patents Form 1/77

Enter the number of sheets for any of the following items you are filing with this form. Do not count copies of the same document

Continuation sheets of this form

Claim(s) 2

Abstract

Drawing(s)

Description

9/4

if you are also filing any of the following, state how many against each item.

Priority documents

Translations of priority documents

Statement of inventorship and right to grant of a patent (Patents Form 7/77)

Request for preliminary examination and search (Patents Form 9/77)

Request for substantive examination (Patents Form 10/77)

Any other documents (please specify)

11.

I/We request the grant of a patent on the basis of this application.

Signature

Marks & Clerk

Date: 16 May 2003

12. Name and daytime telephone number of person to contact in the United Kingdom

M&C Folio: GBP290013

### Activity-directed in vitro protein cloning and identification

5

The present invention relates to cell-free methods for cloning proteins. In particular it relates to a novel cell-free procedure for cloning active proteins from cDNA libraries.

#### 1. Background

10

15

There has been an increasing interest in the development of methods for selection and identification of proteins with defined activities from a large library. With whole genome sequences available, proteomic studies require high throughput methods to assign functions to the proteins encoded and to isolate proteins with desired function. Protein therapeutics and drug design also require powerful selection methods to screen a large combinatorial library or large numbers of combinatorial libraries to improve the function of proteins and to create novel activities (Wang and Saven, 2002, Nucleic Acids. Res. 21, e120)

A number of display technologies have been developed for selection of proteins. 20 Using the principle of coupling phenotype (protein) to genotype (gene), proteins have been successfully displayed on phage, cell surface and virus or ribosome, plasmid and mRNA. Display technology recovers DNA through the functionality of the encoded protein (He and Taussig, 2002, Briefings in functional genomics and proteomics 1, 204-25 212),

30

Ribosome display is a cell-free protein display method which allows the selection and evolution of proteins in vitro (He and Taussig (1997) Nucleic Acids Res. 25: 5153-5134; Hanes and Pluckthun, (1997) Proc. Natl. Acad. Sci. USA 94: 4937-4942). The method generates a library of stable protein-ribosome-mRNA (PRM) complexes from a diversity of DNA molecules by cell-free expression, followed by capture of specific

PRM complexes with a ligand through binding interaction of the displayed nascent protein. The associated mRNA is then retrieved and amplified as cDNA by RT-PCR.

Both prokaryotic and eukaryotic ribosome display systems have been developed and used for selection of peptides, single-chain antibodies, enzymes, stable protein scaffolds and other ligand-binding domains (He and Taussig 2002). A review of tibosome display technology is provided by He & Taussig (2002) Briefings in Functional Genomics and Proteomics, 1(2): 204–212.

A key step in ribosome display is the efficient recovery of genetic material from PRM 10 complexes after selection. Currently, two principal recovery methods are employed. One is a ribosome disruption procedure used in prokaryotic ribosome display, which releases mRNA by the dissociation of ribosome complexes with EDTA followed by RT-PCR (Hanes and Pluckthun, 1997). The other method is an in situ RT-PCR method used in eukaryotic ribosome display (He and Taussig, 1997), which recovers 15 DNA directly from PRM complexes without ribosome disruption through the use of a primer hybridising at least 60 nucleotides upstream of the 3'end of the mRNA in order to avoid the region occupied by stalling ribosome. It has been demonstrated that the in situ RT-PCR procedure is more effective for recovery of DNA from eukaryotic ribosome complexes than the prokaryotic ribosome disruption method (He and 20 Taussig, 2002 Briefings Func Genomics & Proteomics 1: 204-212). This has been used to display antibody fragments as "ARM" (antibody-ribosome-mRNA) complexes. Both methods require a sensitive procedure to recover cDNA from mRNA

25

30

Ribosome display offers advantages over cell-based systems in that it displays a larger library and is capable of continuously introducing new diversities into the selected candidates during selection (He and Taussig 2002). It can also be used to display toxic or unstable proteins or those containing non-natural amino acids including chemically modified amino acids (He and Taussig 2002).

Following ribosome display, the recovered cDNA must be cloned for functional screening of individual proteins. Conventionally, this has been carried out by *E.coli* 

30

cloning and expression, which is a time-consuming and inefficient procedure due to material loss during plasmid construction and transformation. Moreover, PCR amplification of the cDNA pool generates many identical *E.coli* clones, leading to a tedious process of clone screening. Usually, only a fraction of the cDNA population can be recovered and identified by *E.coli* cloning.

The Polymerase Chain Reaction (PCR) can be used to amplify DNA sequences through repeated cycles of template denaturation, primer annealing and elongation. Reverse transcription (RT) coupled with PCR (RT-PCR) combines cDNA synthesis from mRNA templates with PCR amplification to provide a rapid and sensitive method for detection, conversion and recovery of mRNA as DNA. The RT-PCR process can be performed in either one-tube or two-tube formats. In one-tube RT-PCR, the RT and PCR take place successively in a single tube using a mutual buffer for both reactions. In two-tube RT-PCR, RT and PCR are carried out separately. It has been shown that one-tube RT-PCR has the greater sensitivity and that as little as 100 copies of mRNA can be amplified (#TB220, Promega). In two-tube RT-PCR, each step can be optimised separately and it may produce higher yields of DNA in some circumstances (TechNotes 9[6], Ambion).

High sensitivity (i.e. obtaining sufficient amount of DNA from as little template as possible) and high specificity (i.e. amplifying only the desired template) are key to successful PCR. They are affected by many factors including the choice of appropriate DNA polymerases, design of suitable primers, suitable buffers, thermal cycling parameters and also the quality of templates.

A single primer PCR approach has been developed for cloning of unknown DNA sequences (Hermann et al., (2000) BioTechniques 29: 1176-1180) and elimination of primer-dimer accumulation in PCR (Brownie, et al., (1997) Nucleic Acids Res. 25: 3235-3241). The single primer PCR uses one primer in the PCR mixture to amplify DNA having identical flanking sequences at both ends.

Recently, a method termed single-molecule PCR (SM-PCR) has been described which amplifies single molecules of double-stranded DNA through 80 PCR cycles by single

primer PCR (Rungpragayphan et al., 2002. J. Mol. Biol. 318: 395-405). In this method, double-stranded DNA is first amplified by PCR using two primers which introduce a tag sequence flanking both ends of the molecule, after which the modified DNA is diluted into single molecules and used as the template for a second round of PCR amplification using the single tag-specific primer (Rungpragayphan et al., 2002). This does not provide a method for recovering single-stranded cDNA molecules. Conventional PCR by two primers has also been described to amplify single cDNA molecules, but it requires repeated rounds of amplification, with changes of primers in each round (Jena et al., 2000. Mol. Immunol. 37: 265-272). However, none of these methods provided a sensitive procedure for cDNA recovery from mRNA

5

10

15

20

25

For screening protein activity, the cloned DNA must be expressed. Recombinant proteins are usually produced in heterologous hosts, which, in addition to being a time-consuming process, often leads to production of inactive molecules (Stevens et al., 2000. Struct. fold. Des. 8: R177-R185). Cell-free systems, in contrast, can synthesize functional proteins rapidly without the need for cloning in organisms such as E.coli. PCR fragments have been directly used for cell-free synthesis of proteins in an array format for functional protein screening (He and Taussig, 2001 Nucleic acids Res. 29: e73; Rungpragayphan, et al., 2003, FEBS Letters, 540, 147-150). Cell-free systems can generate proteins that heterologous systems cannot produce (www.promega.com/techserv/tntbib.html). Recent improvements have increased protein yield to a level comparable or superior to cell-based expression systems (Madin, K et al., 2000, Proc. Natl. Acad. Sci. USA 97: 559-564). Active proteins at concentrations of Img/ml and 4mg/ml have been produced in E.coli S30 extract and wheat germ system, respectively (Roche RTS100, E.coli HY kit; Madin, K et al., 2000; Sawasaki, T et al., 2002 Proc. Natl. Acad. Sci. USA 99: 14652-14657). Currently, cell-free systems are being used to synthesize proteins for structural and highthroughput proteomics (Sawasaki, T et al., 2002).

30 Using a cell-free system, a procedure termed In Vitro Expression Cloning (IVEC) has been described (Lustig, KD et al., 1997 Meth. Enzymol 283: 83-99), which identifies cDNA clones by cell-free expression of pooled plasmids purified from 50-100 E.coli

clones followed by assays for protein activity. Positive pools are then subdivided until a single plasmid encoding the desired protein is obtained.

Proteomics approaches aimed at screening for protein function have been aided by the use of array technology. Protein arrays are produced by immobilising many hundreds of individual proteins in a defined pattern onto a solid surface (Zhu and Snyder, 2003, Curr. Opin. Chem. Biol. 7: 55-63; Pandey and Mann. 2000, Nature 405: 837-846). In the array format, large numbers of proteins are analysed simultaneously in parallel, providing valuable information on protein function, interaction and expression profiling (Zhu and Snyder. 2003; Pandey and Mann. 2000). Currently, the main limitation to protein array technology is the production of the huge diversity of proteins that form the array elements. To overcome these problems, a cell-free method has been developed, termed DiscernArray<sup>TM</sup>, which creates functional protein arrays directly from PCR-generated DNA by in vitro synthesis of individual tagged proteins on tag-binding surfaces. Tagged proteins are immobilised in situ as they are synthesised. This method is described in PCT Application No. PCT/GB01/03657.

DiscernArray<sup>TM</sup> avoids cloning and *E.coli* expression, providing a rapid route for arraying proteins or domains for which DNA clones are not available. It is also particularly useful for proteins which can not be functionally produced in heterologous hosts. It can be adapted for high-throughput application and automation. This technology has been used to generate arrays of different proteins and protein fragments and demonstrated their use for rapid functional analysis (He and Taussig, 2001).

#### 25

5.

10

15

20

## 2. Brief description of the invention

The present invention provides a method for activity-directed in vitro protein cloning.

30 The present invention provides a cell-free procedure to isolate and identify proteins with defined function from cDNA libraries.

The present invention provides a cell-free method for activity-directed *in vitro* protein cloning comprising cell-free transcription of cDNA and cell-free translation of the resulting mRNA, to produce complexes in which mRNA and the protein encoded thereby are associated, selection of mRNA by protein-directed selection and *in vitro* cloning by conversion of mRNA to cDNA clones. Optionally, cell-free expression of cloned DNA may be used to produce proteins, which may, optionally, be assayed to determine function.

A method for cell-free selection and identification of proteins from DNA or mRNA libraries, said method comprising:

- (a) Selection of proteins through ligand-binding by cell-free protein display;
  - (b) Recovering nucleic acids by cDNA synthesis;
- (c) Separation of the cDNA into a small group or single molecules, e.g. by
   serial dilution;
  - (d) In vitro cloning by PCR amplification of the cDNA dilution;
  - (e) Protein production and identification by cell-free expression of the cloned PCR fragments.

## 20 A method comprising:

- (a) Display of an mRNA-protein complex,
- (b) Selection of mRNA-protein complex(es) of interest,
- (c) Reverse transcription of mRNA to cDNA,
- (d) Conversion of cDNA to a double stranded cDNA clone, preferably by

#### 25 PCR.

5

- (f) Optionally repeating (a) to (d),
- (g) Optionally translating cDNA into protein,
- (h) Optionally assaying the protein.
- The mRNA-protein complex is preferably produced from cDNA, for example from a PCR library. Starting with PCR libraries, cDNA encoding desired proteins are first enriched, for example by ribosome display through ligand-binding in which cDNA from the PCR library is subjected to cell free translation to produce protein-tibosome-

mRNA complexes, in which the nascent protein and its encoding mRNA are associated. Selection for mRNA is achieved by selecting for the associated protein. The mRNA is then converted to cDNA, which is cloned *in vitro* by PCR to generate individual PCR fragments, which can be used for cell-free expression to identify proteins with desired activity. The entire process can be automated, making it possible to simultaneously screen many different libraries (> 100 libraries) or a very large library (>10<sup>14</sup> members) against various ligands. This unique capability allows mutant libraries spanning an entire region of a protein to be screened, offering a powerful tool for rapid evolution of proteins *in vitro*.

10

15

20

25

30

5

The protein-mRNA complex can be produced using cell-free systems such as cell free display. Transcription of the cDNA to mRNA and translation of the mRNA to protein can be coupled (performed together) or uncoupled (performed separately). Cell free display methods include mRNA display (Roberts R.W. & Szostak J.W. (1997) Proc. Nat. Acad. Sci. USA, 94, 12297 – 12302) and ribosome display. In a preferred method, cell-free translation of cDNA to produce complexes in which mRNA and the protein encoded thereby are associated is performed by ribosome display. Ribosome display is carried out fully in vitro, which overcomes some of the limitations of cell-based display systems. Ribosome display systems such as, E.coli S30 or rabbit reticulocyte can be used to produce protein-ribosome-mRNA (PRM) complexes in which the individual nascent proteins are linked with their corresponding mRNA molecules.

PRM complexes are produced by ribosome stalling such that the nascent protein and mRNA remain associated. Strategies to achieve this include the addition of antibiotics such as rifampicin or chloramphenical (for prokaryotic ribosomes), or such as cyclohexamide (for eukaryotic ribsomes) to halt translation at random. Alternatively, the ribosome may be caused to stall at the 3' end of the mRNA template because of deletion of the stop codon from the mRNA template; the stop codon normally being recognised by release factors that trigger detachment of the nascent polypeptide.

In general, ribosome display constructs should contain a promoter (T7, SP6 or T3) and a translation initiation signal such as a Shine-Dalgarno (prokaryotic) or Kozak

ווטער בטואבטוא

(eukaryotic) sequence. A consensus sequence for protein initiation in both *E. coli* and eukaryotic systems has also been described. To enable the complete nascent protein to be displayed and fold into its active conformation, a spacer domain of at least 23-30 amino acids length is required at the C terminus, to allow the protein to exit completely from the ribosome 'tunnel'. The spacer also provides a known sequence for the design of primers for RT-PCR recovery. A number of different spacers have been successfully used, including the constant region of immunoglobulin kappa chain (Ck), gene III of filamentous phage M13 and the Ch3 domain of human IgM. Spacer length has been shown to affect display efficiency: a spacer of 116 amino acids was more efficient in displaying proteins than its shorter partners. To remove the stop codon from DNA, a 3' primer lacking the stop codon is used during PCR construction. Constructs designed for prokaryotic *E. coli* display should incorporate sequences containing stem-loop structures at the 5' and 3' ends of the DNA to stabilise mRNA against degradation by RNase activities in *E. coli* cell-free systems.

15

20

25

30

5

10

The displayed polypeptide/protein can be a full length protein, or a fragment thereof, a protein domain, or a polypeptide, and can be encoded by synthetic DNA.

Display systems allow selection of genetic material through the functional properties of the protein, usually as a binding reaction. Selection of an mRNA-protein complex can be achieved by selection of the protein moiety of interest, e.g. by ligand binding. The displayed proteins can be captured through interaction with one or more ligand(s). The ligand may be an antibody, or a fragment thereof, or an antigen. The ligand(s) may be coupled to a solid support such as a well, tube or bead; the coupling may be covalent or non-covalent. The mRNA-protein complex may be selected via interaction with soluble biotinylated proteins followed by capture on streptavidin-coated surfaces. It is essential to keep the PRM complexes intact during selection. This is achieved by maintaining the complexes at 4°C with elevated magnesium concentrations, e.g. 5mM for eukaryotic ribosome complexes, and 50 mM for E. coli complexes. Under such conditions, complexes can be stable for at least two weeks. To obtain effective selection, background binding needs to be eliminated, and various blocking reagents have been applied to reduce non-specific sticking. A defined

20

25

selection condition or strategy often results in the production of proteins with desired properties.

The *in vitro* PCR cloning may be performed by separate RT and PCR reactions or a combined RT-PCR reaction, preferably using a single molecule or group of several molecules as template(s).

The substrate for reverse transcription may be mRNA dissociated from protein or mRNA bound in complexes. Release of mRNA from ribosome complexes can be achieved in a ribosome disruption procedure using EDTA. Recovery of DNA from PRM complexes without ribosome disruption can be performed using an *in situ* RT or RT-PCR method.

A denaturation step, using chemical or heat denaturation (e.g. 65° C for 5 minutes)

may be performed prior to the reverse transcription (RT), or RT-PCR in vitro cloning procedures.

Reverse transcription is preferably carried out using an RT primer to introduce a flanking sequence into single-strand cDNA as a template for the subsequent amplification to be carried out by single primer PCR. This differs substantially from previous single primer PCR methods. Earlier methods did not indicate how DNA would be recovered from mRNA. In contrast, the present invention uses mRNA as the template. The flanking sequence can be designed from the 5' region of the mRNA rather than an artificially designed sequence and introduced into single stranded cDNA rather than double stranded DNA.

According to the present invention, a preferred RT primer for use with mRNA comprises a flanking sequence and a tail hybridising sequence.

The flanking sequence is designed from a knowledge of the 5' region of the mRNA, and is a sequence which will hybridise with DNA complementary to a part of the mRNA 5' region. Typically the flanking region is identical to a part of the mRNA 5'

region. The flanking region is usually in the range of an 8-mer to a 30-mer, such as a 10-mer to a 20-mer, preferably around a 15-mer.

The tail hybridising sequence is designed from a knowledge of the 3' region of the mRNA and hybridises to a part of the mRNA 3' region. Typically the tail hybridising sequence is identical to a sequence complementary to a part of the mRNA 3' region. The tail hybridising sequence region is usually in the range of an 8-mer to a 30-mer, such as a 10-mer to a 20-mer, preferably around a 15-mer. In the primer, the tail hybridising sequence is 3' relative to the flanking sequence.

10

15

20

25

5

The RT-PCT procedure may be carried out in one tube or two tube format, two tube format is preferred.

Conversion of cDNA to duplex cDNA clone, is preferably carried out by PCR, most preferably using a single primer e.g. by a method known in the art or by a novel method such as that described herein.

In vitro cloning, i.e. conversion of mRNA to cDNA, is preferably performed using a novel ultra-efficient RT-PCR procedure described herein, which is capable of recovering a single cDNA molecule. This ultra-sensitive RT-PCR procedure allows cDNA to be cloned in vitro by PCR after ribosome display, avoiding laborious and inefficient E.coli cloning. In this procedure, a novel primer has been designed to produce single-stranded cDNA with a flanking sequence at both ends, allowing the generated cDNA to be amplified by subsequent PCR using a single primer. Single primer PCR avoids the need to optimise primer ratios and annealing temperatures, thus increasing PCR sensitivity and specificity with a wide choice of DNA polymerases. By using this highly sensitive recovery method, rare species can be isolated from very large libraries.

Dilution of nucleic acid molecules may be performed to provide a single, or small number of nucleic and molecules in each reaction. Thus, mRNA may be diluted prior to the RT reaction, and/or cDNA may be diluted prior to the PCR reaction to provide a single or a small number e.g. 1, 2, 3, 4, 5 or 6, nucleic acid molecules per reaction.

The single nucleic acid molecules and/or small groups of nucleic acid molecules can be produced by serial dilutions.

The nucleic acid molecule(s) can be double-stranded DNA, single-stranded DNA (such as cDNA), mRNA and their associated complexes such as protein-ribosome-mRNA complex(es) and mRNA or DNA-protein fusion.

The RT, PCR or RT-PCR may be carried out using one or more primer(s) designed for recovering the target template(s).

10

15

20

25

30

5

One particular application of this novel RT-PCR procedure is recovery of cDNA from protein display complexes, such as ribosome display complexes. In this application, RT and PCR reactions are carried out separately in a two-tube format and each step is optimised to improve the sensitivity and specificity of DNA recovery. We found that a pre-heat step at 65°C was required prior to an in situ RT process to ensure full-length cDNA synthesis. The design of the flanking sequence is based on a consensus sequence present in the 5' region of the mRNA population (Figure 2). In the example described here, it is the sequence occurring from the start site of transcription to the translation initiation site (ATG). 5'-GAACAGACCACCATG-3. The hybridising sequence is based on a consensus sequence located at mRNA 3' region (Figure 2). Reverse transcription using the novel RT primer, comprising both the flanking sequence and the hybridising sequence in accordance with the invention, incorporates the flanking sequence into the single-stranded cDNA at both ends, allowing the resultant cDNA to be amplified by a single primer composed of the flanking sequence (Figure 2). Single primer PCR avoids the need to optimise primer ratios and annealing temperatures, thus increasing PCR sensitivity and specificity with a wide choice of DNA polymerases. For recovering DNA such as from ribosome complexes, the introduction of the flanking sequence to the cDNA also distinguishes it from the original input DNA, eliminating any potential PCR amplification of the input template. Following this highly sensitive RT-PCR recovery, a second PCR is easily carried out using two primers to recreate a full-length DNA construct (Figure 3). Such a construct obtained from cDNA of a ribosome display complex is suitable for subsequent cycles of ribosome display.

10

15

20

Ribosome display can be used in methods of the invention as a powerful means for bringing about protein evolution in vitro, since mutation can be continuously introduced into DNA pools in subsequent cycles. Successive DNA diversification, followed by ribosome display/selection under defined selection conditions, can be used to produce mutant proteins, e.g. antibody mutants with improved affinity and stability. By using a low fidelity DNA polymerase for DNA amplification, ribosome display has selected antibody mutants with up to a 40-fold improvement in affinity over the original antibody fragment. This shows that protein evolution occurs during ribosome display cycles. Various PCR strategies, such as error-prone PCR, DNA shuffling or a combination of site-directed mutagenesis, error-prone PCR and H-CDR3 shuffling (authors' unpublished data), can be used to increase mutation rate, leading to the selection of evolved antibody fragments. Mutation can also be accumulated in mRNA during display cycles by the inclusion of Qβ RNA-dependent RNA polymerase in the translation mixture.

The novel design of an RT primer which incorporates a flanking sequence into single-stranded cDNA at both ends e.g. after ribosome display, makes it possible to amplify the resultant cDNA by single primer PCR in a sensitive, specific and reproducible way. The cDNA with a flanking sequence can be amplified up to 100 cycles without producing non-specific DNA bands or smearing. Thus, a single cDNA molecule can be efficiently amplified through 65 PCR cycles. To our knowledge, this is the first report of such a sensitive RT-PCR method for recovery of cDNA.

We have shown that cDNA can be amplified up to 100 cycles without producing nonspecific DNA bands or smearing. This technology makes it possible to clone cDNA in
vitro by PCR amplification of individual cDNA molecules following appropriate
dilution. The PCR clones can be stored or converted into a suitable construct for
downstream processes such as protein expression in cell-free or cell-based systems.

PCR assembly may be used to generate a DNA construct for expression of protein
encoded by the cloned cDNA. The cDNA may be introduced into a host cell for
expression, or cell-free expression may be used to produce protein. The in vitro
cloning offers a number of advantages over E. coli cloning in that (i) it is rapid,

avoiding repeated screening of identical *E.coli* clones and (ii) it is more efficient, allowing clones covering an entire cDNA population to be recovered and screened.

Since this invention is capable of amplifying a single cDNA molecule, it is also possible to clone cDNA by PCR through cell-free display of proteins without using E. coli cells, providing a route to obtaining PCR clones through the encoded protein. This in vitro PCR cloning would offer a number of advantages over E. coli cloning strategy in that (i) cDNA molecules are separated by dilution before they are amplified, so that individual cDNA species can be rapidly obtained by PCR, avoiding the time-consuming and laborious procedure of identifying E. coli clones after DNA amplification and cell transformation and (ii) it is possible to obtain individual DNA clones covering an entire cDNA population by high throughput PCR, whereas E. coli cloning generally only recovers a fraction of the population due to DNA manipulation and cell transformation, procedures resulting in loss of material.

15

5

10

The resultant PCR clones can be directly used for cell-free expression to identify desired candidates (Figure 1).

The cell-free expression system may utilise *E.coli* S30 extract, wheat germ system and/or rabbit reticulocyte lysate in either coupled or un-coupled manner.

The proteins (individual or a group) produced in cell-free expression may be synthesised either in a soluble form or immobilised on a solid surface, for example through a tagged sequence on a solid surface in an array format.

25

Assays may be performed to investigate the function of the protein. The protein activity screening includes, but is not limited to, assays for ligand-binding and enzyme activities.

The invention also provides a method comprising: (i) selection based on protein function by ribosome display to enrich the candidates from a large population; (ii) PCR cloning of the selected cDNA; (iii) cell-free expression of the PCR clones for functional screening. This invention combines protein selection, an in vitro cloning

strategy and protein screening, providing a more effective and high-throughput tool than existing methods for selection and identification of proteins. In addition, it allows the entire population of the selected cDNA to be potentially cloned and functionally screened.

5

15

20

25

30

The whole or parts of a methods of the invention may be carried out by automation using a robot.

Using the novel method, it is possible to automate protein display, offering a powerful tool for simultaneously screening multi-display libraries or selection of different binders to various antigens.

Furthermore, the full process can be automated, offering a high-throughput format for simultaneously screening large numbers of combinatorial libraries against various ligands. This unique capability allows much larger library sizes to be screened. It also provides a powerful tool to screen mutant libraries spanning the entire region of a protein for rapid evolution of proteins *in vitro*.

The present invention also provides kits for use in a method of the invention. The kit comprises a supply of components for use in a method of this invention, and typically comprises a supply of primer designed according to the invention, and one or more of a supply of dNTP, a supply of reverse transcriptase, a supply of ribonuclease inhibitor, buffer, RNase-free water. The kit might also include PCR components. The kits of this invention suitably include a set of instructions for use of the components in accordance with a method of this invention.

## 3. Detailed description of procedures of the invention:

#### 3. 1. Materials

- (1) TNT T7 quick for PCR DNA (Promega; Cat. #L5540).
- (2) RNase-free DNase I (Roche; Cat. #776 785).
- (3) Taq DNA polymerase (Qiagen; Cat. #201203).

- (4) SUPERase In (Ambion)
- (5) ThermoScript (Invitrogen)
- (6) Phosphate-buffered saline (PBS): 1.63mM Na<sub>2</sub>HPO<sub>4</sub>, 1.47mM KH<sub>2</sub>PO<sub>4</sub>, 2.68mM KCl and 136mM NaCl, pH 7.4.
- 5 (7) 10x DNase I buffer (400mM Tris-HCl, pH 7.5; 60mM MgCl<sub>2</sub>; 100 mM NaCl and 10mM MnCl<sub>2</sub>).
  - (8) 2 x dilution buffer: PBS containing 10 mM magnesium acetate or 4mM glutathione solution containing 10mM magnesium acetate, GSSG: GSH=1:1), stored at 4°C.
- (9) Washing buffer: PBS containing 0.1% Tween 20 and 5 mM magnesium acetate, stored at 4°C.
  - (10) Reverse transcription mix 1: 1.3uM oligo primer and 1.6 mM dNTPs
  - (11) Reverse transcription mix 2: 1x cDNA synthesis buffer, 1.25mM DTT, 20U SUPERase In and 15U ThermoScript

15

#### 3.2. Methods

A method of the invention comprises:

#### A. Ribosome display

- 20 (1) Generation of protein-ribosome-mRNA (PRM) complexes by ribosome display as described (He et al., 1999, J. Immunol. Methods 105-117)
  - (2) Exposure of the PRM complexes to a ligand which is coated onto wells of a microtitre plate (or linked to beads);
  - (3) Incubation for PRM-ligand interaction;
- 25 (4) Washes
  - (5) Reverse transcription to generate cDNA as described. The resultant cDNA can be amplified either as a pooled template for further ribosome display cycles (repeat step1-5), or continue as follows:

#### 30 B. In vitro PCR cloning

 Dilution of the cDNA successively to a point where it contains either a small group of molecules, a single molecule or no molecule (Jena, et al., (2000) Molecular Immunology 37, 265-272);

- (2) PCR amplification to generate individual fragments or a small group of fragments;
- (3) PCR assembly to generate DNA construct suitable for cell-free expression;

## 5 C. Cell free expression of proteins into arrays format for functional screening

- (1) Cell-free synthesis of proteins either as a soluble form or immobilised onto a solid surface in the format of protein arrays.
- (2) Assay for protein function.

## 10 3.3. A procedure for automation of the method:

#### A. Ribosome display

15

- (2) Using a robot such as Tecan miniprep 75, add to each well of a microtitre plate 10ul of PCR library, followed by 40ul of TNT mixture containing 0.02mM methionine and 1mM Mg acetate;
- (3) Incubation at 30°C for 1 hr;
- (4) Adding to each well 20ul mixture containing 7 μl of 10x DNase I buffer and 100U DNase I;
- (5) Incubation at 30°C for 20 min;
- 20 (6) Adding to each well 70 μl of cold 2x dilution buffer;
  - (7) Transferring the above mixture (140ul) to well of Nucleolink<sup>TM</sup> strips (NUNC) coated with ligand(s);
  - (8) Incubation at 4°C for 2 hrs with gentle shaking;
  - (9) Washes 5 times by the washing buffer, then two times by sterilised H<sub>2</sub>O;
- 25 (10) Adding to each well 12ul of reverse transcription mix 1 and incubation at 65°C for 5 min, followed on ice for 1 min.
  - (11) Adding to each well 8ul of mix 2 and incubation at 60°C for 45 min followed by 85°C for 5 min.
- (12) PCR amplification of the resultant cDNA either as a pooled template for further ribosome display cycles (repeat step1-5), or as individual molecules as follows:

#### B. In vitro PCR cloning

- (1) Division of cDNA molecules by serial dilution and dispersion into microtitre wells.
- 5 (2) Amplification of the cDNA dilution by adding a single primer PCR mixture as described. Thermo-cycling on PCT200 for 35-65 cycles
  - (3) Agarose gel analysis;
  - (4) PCR assembly by mixing individual PCR fragments with a plasmid containing essential elements for cell free expression (He and Taussig, 2001). Thermo-cycling as in the step 2;
  - (5) Agarose gel analysis;

10

20

25

## C. Cell-free expression for functional assay

- (1) Adding the individual PCR constructs into microtitre wells containing 25ul cell-free extract (such as E.coli S30 extract, Rabbit reticulocyte lysate or wheat germ). For the generation of immobilised proteins, the microtitre wells are pre-coated by a tag-binding reagent (He and Taussig, 2003, J. Immuno. Methods. 274, 265-270; PCT Patent Application No. PCT/GB01/03657).
  - (2) Activity assay based on the function of the proteins to be analysed.

Depending on the type of a robot to be used, the above automation procedure could be used for either semi-automation using a liquid handling robot (eg. Tecan Miniprep 75) or for full automation (eg. Using Tecan TRAC system with RoMa robot arms to move plates to an incubator, PCR blocks and a plate washer).

## 3. 4 Detailed methods for the RT-PCR Reaction:

#### 3.4.1 Materials

- 30 Recovery from ribosome display is used as an example:
  - (1) Primers:
    - (a) HuRT: 5'-GAACAGACCACCATGACTTCGCAGGCGTAGAC-3' is used for reverse transcription. The flanking sequence is underlined

(b) Kz1: 5'-GAACAGACCACCATG-3' is the flanking sequence and used for single primer PCR

The following primers (c & d) are used for re-creation of the original full-length DNA construct for subsequent ribosome display cycles

(c)T7Ab:

5'GGATCCTAATACGACTCACTATAGGGAACAGACCACCATG(C/G)AGGT(G/C)CA(G/C)CTCGAG(C/G)AGTCTGG -3'

The T7 promoter is indicated in bold.

10 (d)Hurex-Ck:

5

5'-CTCTAGAACACTCTCCCCTGTTGAAGCTCTTTGTGACGGGC GAGCTCAGGCCCTGATGGGTGACTTCGCAGGCGTAGACTTTG-3'

- (2). Enzymes:
- 15 (a) ThermoScript (Invitrogen)
  - (b) SUPERase In (Ambion)
  - (c) Taq DNA polymerases (Qiagen)
  - (d) SuperScript<sup>TM</sup> II Rnase H Reverse transcriptase (Invitrogen)

## 20 3.4.2 Method

25

Figure 3 shows the principle of the RT-PCR for DNA recovery from ribosome complexes. RT is first carried out on intact ribosome complexes using the primer HuRT (see 3.4.1). The resultant cDNA with the flanking sequence is then amplified by PCR using the primer Kz1.

## 3.4.3 Details of the procedure

Ribosome-bound mRNA generated by ribosome display in the form of PRM complex is usually captured by a ligand-coated tubes or magnetic beads (He et al., (1999) J. Immuno. Methods. 231: 105-117) and used for DNA recovery by the following procedure.

10

30

## 3.4.3.1. Reverse Transcription

- (1) To each sample, 12  $\mu$ l of mixture containing 1 $\mu$ l of 16 $\mu$ M HuRT and 2 $\mu$ l of 10 $\mu$ M dNTP is added. The mixture is heated at 65°C for 5 min, and then quickly placed on ice for at least 30 sec.
- (2) A reverse transcriptase is used to synthesise cDNA according to the manufacturer's instructions. The followings are two examples:
  - (a) For Thermoscript, 8μl of mixture containing 4μl of 5 x cDNA synthesis buffer (included in the kit), 1μl of 100mM DTT, 1μl of SUPERase In (20U), 1 μl of Thermoscript (15U) is added to the mixture from step 1 above and the incubation is carried out at 60°C on a thermocycler for 45 min followed by 5min at 85°C. Finally on hold at 10°C
- (b) For SuperScript, 8ul of mixture containing 4ul of 5 x First-Strand buffer (included in the kit), 1ul of 100mM DTT, 1ul of SUPERase In (20U), 1 ul of Superscript<sup>TM</sup> (200U) is added. The mixture is incubated at 46°C for 45 min followed by 5 min at 85°C.

The cDNA mixture produced is then transferred to a fresh tube for subsequent 20 amplification by single primer PCR

## 3.4.3.2. Single Primer PCR of single stranded cDNA

Single primer PCR is carried out using a DNA polymerase according to manufacturer's instructions. The following PCR mixture is an example using Qiagen Taq polymerase.

10x buffer 5ul
5xQ buffer 10ul

dNTPs (2.5mM) 4ul
Primer Kzl (8uM) 3ul
Taq polymerase 0.5ul (2.5units)
cDNA 1-2ul

cDNA 1-2ul  $H_2O$  to 50ul

Thermal cycling is carried out as follows:

35 cycles of 94°C (1min), 48°C (1min) and 72°C (1.2 min), followed by one cycle at 72°C for 8 min and finally placed on hold at 10°C.

5 The single primer PCR product is analysed and extracted from 1% agarose gel if necessary.

## 3.4.3.3 Conversion of recovered DNA into a suitable construct for subsequent manipulations

10

15

The single primer PCR product can be easily converted into a suitable form for further manipulations such as protein expression in a cell-free system or in *E.coli*, or repeated cycles of ribosome display by a second PCR to introduce some necessary elements (e.g T7 promoter, restriction enzyme sites, purification tag etc) into the DNA construct. The following is an example of re-creating the full-length DNA construct for continuous ribosome display (see Figure 3). The primers used are T7Ab and Hurex-Ck (Materials 3.4.1), which contain T7 promoter and a part of the constant region for the extension of the construct into its original full-length (Figure 3).

## 20 The second PCR is carried out as follows:

	10x buffer		5ul
25	5 x Q buffer		10ul
	dNTPs (2.5mM)		4ul
	T7Ab (16uM)		1.5ul
	Hurex-Ck (16uM)		1.5ul
	Gel-purified DNA fragment		50-100ng
	Taq polymerase (Qiagen)		0.5ul (2.5units)
	H <sub>2</sub> O	to	50ul

30

Thermal cycling is carried out by 20-25 cycles of 94°C (1min), 60°C (1min) and 72°C (1.2 min), followed by 72°C for 8 min and finally placed on hold at 10°C.

After agarose gel analysis, the PCR product is ready for ribosome display.

## List of figures

- Figure 1. Scheme for activity directed in vitro protein cloning and identification.
- Figure 2. DNA construct for the generation of ribosome display complex.
- Figure 3. The principle of the novel RT-PCR for cDNA recovery from ribosome complex.
  - Figure 4. Efficiency of two tube RT-PCR.
  - Figure 5. Time course of single primer PCR.
  - Figure 6. Choice of PCR Enzymes.
- 10 Figure 7. PCR amplification of a single cDNA molecule
  - Figure 8. Effect of pre-heating on DNA recovery efficiency
  - Figure 9. Integrity of ARM complexes

## 4. Examples

15

# 4.1. Selection of an anti-CEA antibody fragment from a human naïve antibody library.

A sheep anti-CEA antibody fragment in the V<sub>H</sub>/K format was mixed with a human naïve antibody library in a ratio of 1:100 (anti-CEA: naïve library). The mixture was subjected to ribosome display and selected by CEA-coated tube. Since the library and anti-CEA contain the constant region (Cκ) of human k chain as the spacer (He and Taussig, 1997), anti-κ antibody coated tube was also used for selection as a control. After one cycle of ribosome display, the resultant cDNA was amplified by the single primer PCR either as a pool or serial dilution (from 1: 10 to 10<sup>7</sup>) by the single primer PCR.

A suitable dilution was chosen for the generation of individual PCR fragments. PCR fingerprinting using restriction enzyme MavI was carried out to identify the PCR fragment. This showed the proportion of anti-CEA fragment. The anti-CEA fragment was then converted into the construct by PCR assembly for cell free expression using E.coli S 30 extract. The anti-CEA was made in either an soluble form or immobilised

through a his-tag sequence onto the microtitre wells pre-coated with nickel ions. The activity of the anti-CEA fragment was analysed by CEA specificity.

## 4.2 The novel RT-PCR procedure is more effective in the two-tube format

5

10

20

25

DNA recovery efficiency was compared using the novel RT-PCR in either two-tube or one-tube formats. Commercially available one-tube RT-PCR kit and two tube RT-PCR systems from Invitrogen were chosen as they use the same enzymes. Antibody-ribosome-mRNA (ARM) complexes captured by carcinoembryonic antigen (CEA) after ribosome display were subjected to RT-PCR using the two systems side by side with identical primers. Analysis of the RT-PCR products showed while one tube RT-PCR produced no DNA, the two-tube system generated very strong DNA recovery (Figure 4), demonstrating that the novel RT-PCR is more effective in two-tube format.

## 4.3 Single primer PCR efficiently and specifically amplifies cDNA with the flanking sequence.

We have tested the sensitivity and specificity of the single primer PCR to amplify cDNA with the flanking sequence by carrying out a time course of 100 PCR cycles. In this experiment, cDNA generated after RT was diluted 10,00 times and PCR was carried out as in total volume of 200ul. During PCR cycles, 20ul was taken at the following time points: 35 cycles, 45 cycles, 55 cycles, 65 cycles, 75 cycles, 85 cycles and 100 cycles. Gel analysis of PCR product showed that a single DNA fragment was specifically amplified up to 100 cycles without detection of DNA smearing or additional products (Figure 5).

## 4.4. Single primer PCR works with a wide choice of DNA polymerases.

A variety of DNA polymerases from different commercial companies were tested for the single primer PCR. Figure 6 showed that most Taq polymerases and proofreading enzymes amplify the resultant cDNA very effectively, indicating it is suitable for a wide range of DNA polymerases without the need for optimisation.

10

15

20

30

## 4.5. Single primer PCR is capable of amplifying single cDNA molecules

To see whether a single cDNA molecule can be recovered by single primer PCR from a cDNA mixture generated by ribosome display, a naïve human antibody library was subjected to ribosome display and BSA selection. After RT, the resultant cDNA was first quantified by PCR comparison with a known amount of template. This gave an indication that a cDNA concentration of 10fg/ul (equivalent to 1000 molecules) was obtained (data not shown). The cDNA was then serially diluted to a point where it contained no more than one molecule/ul. This was tested by PCR of serial dilutions (1:10, 1:10 $^2$ , 1:10 $^3$  and 1:10 $^4$ ). No PCR product was detected in the dilution of 1:10 $^4$  in 80 cycles and only half of samples from 1:103 dilution produced PCR product, suggesting cDNA concentration of no more than one molecule/ul was produced in the 1:10<sup>3</sup> dilution (data not shown). In order to test if individual cDNA molecules could be recovered by PCR from the 1:10<sup>3</sup> dilution, DNA fingerprinting by Mavel restriction and direct DNA sequencing were used. Figure 7a showed while a 1:10 dilution produced a V<sub>H</sub>/K band by 35 PCR cycles, templates from 1:10<sup>2</sup> and 1: 10<sup>3</sup> dilutions required 65 cycles to reveal the band, DNA fingerprinting revealed that whereas PCR fragment from 1:10 and 1:10<sup>2</sup> dilution gave a similar digestion pattern, the templates from 1:103 produced different DNA patterns among the duplicates (Figure 7b), suggesting a single cDNA was amplified from 1:103 dilution. Direct DNA sequencing of the PCR fragments from 1:103 dilution showed only one unique sequence from each sample, further confirming the recovery of individual cDNA molecules by single primer PCR.

## 4.6 Pre-heating ribosome display complexes at 65°C improves the synthesis of full-length cDNA

Pre-heating of ribosome ARM complexes at 65°C prior to RT-PCR improves the recovery of full-length cDNA. Figure 8a shows a comparison of RT-PCR with or without the pre-heating step. In this experiment, ARM complexes after CEA selection were treated either by heating at 65°C for 5 min or without the pre-heating step, followed by cDNA synthesis at 60°C using ThermoScript (Invitrogen). This showed that about 10 fold more V<sub>H</sub>/K DNA was recovered from the pre-heated sample.

10

15

20

24

Samples without pre-heating gave very poor DNA detection even at optimised temperatures (50°C to 60°C) for the activity of ThermoScript (Figure 8a). Increased full-length V<sub>H</sub>/K recovery was also observed by the pre-heating step when a normal SuperScript<sup>TM</sup> II Rnase H Reverse transcriptase (Invitrogen) was used at 46°C (Figure 8b).

## 4.7 The ribosome complex remains intact at 65°C

The integrity of ARM complexes at the elevated temperature was tested by reverse transcription of the ribosome-attached mRNA with primer [GGTGACTTCGCAGGCGTAGACTTTG], which anneals at about 100bp upstream stalled ribosome complex (Figure 9a) and with [ACTCTCCCCTGTTGAAGC] which anneals at the 3' region of mRNA covered by the stalled ribosome. The RT process was carried out as described in the Method (see 3.4.2) using Thermoscript. While little or no PCR products were observed with primer 0, which is blocked in an intact complex (Figure 9b, lane 4), the full-length product was generated with the upstream primer 1 (Figure 9b, lane 7). A PCR template terminating at the primer I site that worked in the first round was then employed (Figure 9a). Again, while the 3' terminal primer 1 did not produce full-length PCR product (Figure 9b lane 11), an upstream primer 2 [TGCTGAGGCTGTAGGTGC] was successful (lane 14). This showed that the 3'end of the mRNA was inaccessible to a primer, probably due to the occupancy of a stalled ribosome under the RT conditions. This suggests that ribosome complexes remained intact after pre-heating at 65°C.

#### CLAIMS:

5

15

25

- 1. A method for cell-free selection and identification of proteins from DNA or mRNA libraries, said method comprising:
  - (a) Selection of proteins through ligand-binding by ribosome display;
  - (b) Recovering nucleic acids by cDNA synthesis;
- (c) Separation of the cDNA into a small group or single molecules by serial dilution;
  - (d) In vitro cloning by PCR amplification of the dilution;
- (e) Protein production and identification by cell-free expression of the cloned PCR fragments.
  - 2. The method as in claim 1, wherein the display method for selection of proteins is any display method using a cell-free system or a modification of these methods.
  - 3. The method as in claims 1-2, wherein the displayed proteins include whole proteins, protein domains and polypeptides encoded by synthetic DNA.
- 20 4. The method as in claims 1-3, wherein the displayed proteins can be captured through interaction with ligand(s).
  - 5. The method as in claims 1-4, the ligand (s) is coupled to wells, tubes or beads either covalently or non-covalently.
  - 6. The method as in claim 1, wherein the *in vitro* PCR cloning is performed by PCR or RT-PCR using a single molecule or group of molecules as template(s).
- 30 7. The method as in claim 5, the PCR or RT-PCR is carried out by a single primer method.

8. The method as in claim 5-6, wherein the template(s) include double-stranded DNA, single-stranded DNA (such as cDNA), mRNA and their associated complexes such as protein-ribosome-mRNA complex(es) and mRNA or DNA-protein fusion.

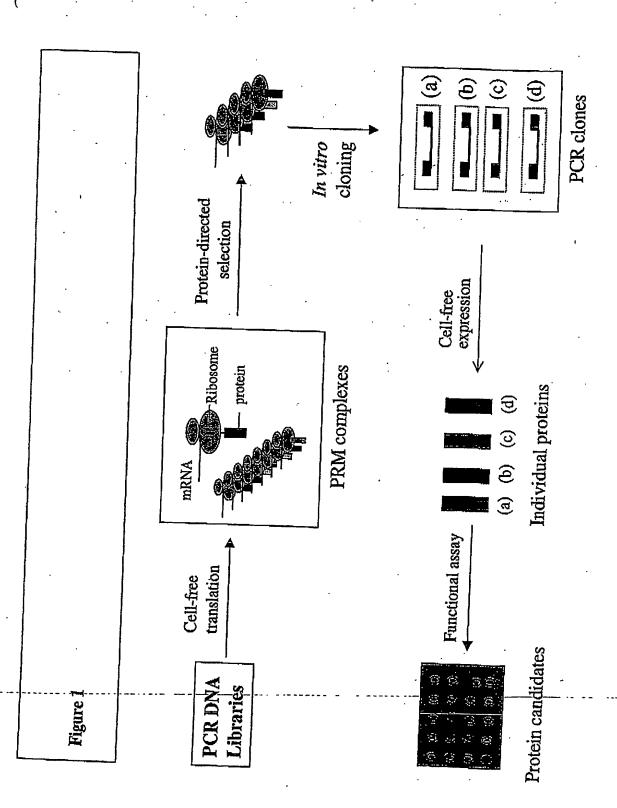
5

- 9. The method as in claims 5-7, wherein the single molecules and group of molecules are produced by, but not limited to, serial dilutions.
- 10. The method as in claim 5, wherein the PCR or RT-PCR is carried out using primer(s) designed for recovering the target template(s)
  - 11. The method as in claim 1, wherein the cell-free expression system includes *E.coli* S30 extract, wheat germ and rabbit reticulocyte lysate in either coupled or un-coupled manner.

15

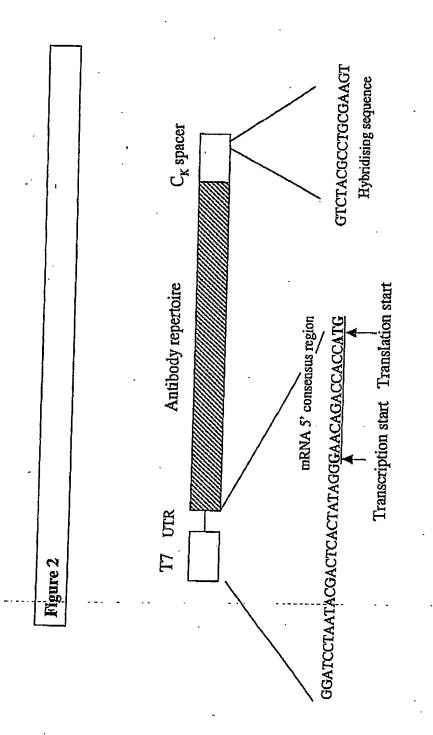
- 12. The method as in claim 11, wherein the proteins (individual or a group) are synthesised either in a soluble form or immobilised through a tagged sequence on a solid surface in an array format.
- 20 13. The method as in claim 1, wherein the protein activity screening includes, but not limited to, assays for ligand-binding and enzyme activities.
  - 14. The method as in claim 1, wherein the whole or parts of the method is carried out by automation using a robot.

25



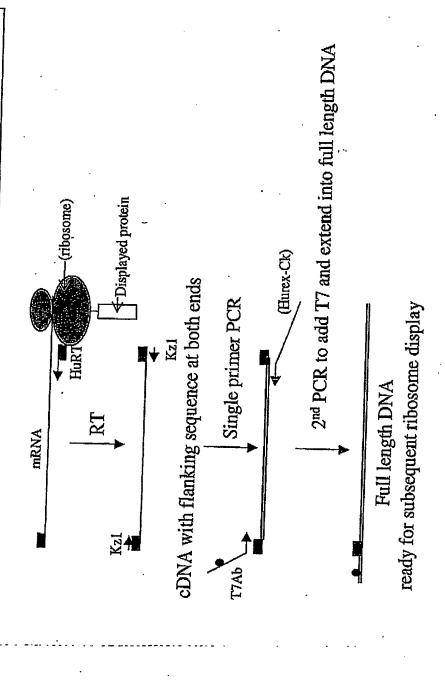
r.J.

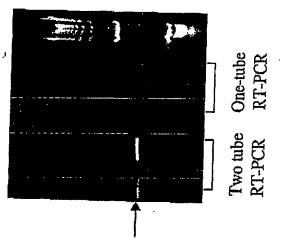
2/9

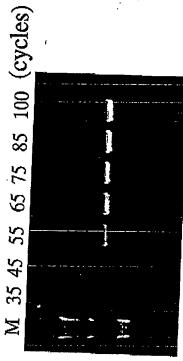


HuRT primer: 5'-GAACAGACCACCATGACTTCGCAGGCGTAGAC-3'
Planking sequence
Hybridising sequence

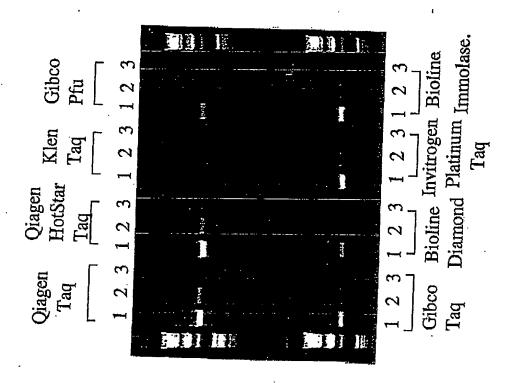
KZ1 primer: 5'-GAACAGACCACCATG-3'

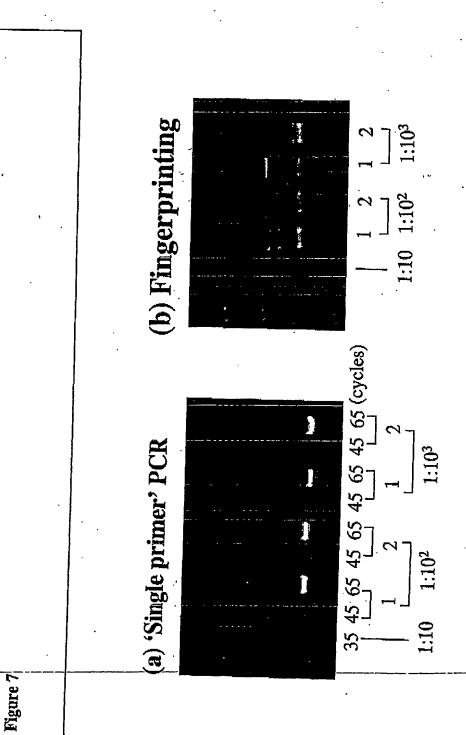


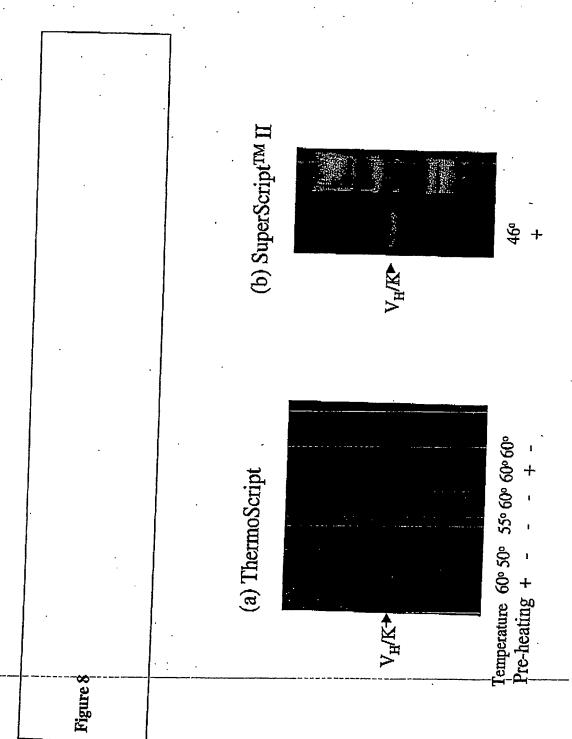












# This Page is Inserted by IFW Indexing and Scanning Operations and is not part of the Official Record

## **BEST AVAILABLE IMAGES**

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images include but are not limited to the items checked:

BLACK BORDERS			
☐ IMAGE CUT OFF AT TOP, BOTTOM OR SIDES			
FADED TEXT OR DRAWING			
BLURRED OR ILLEGIBLE TEXT OR DRAWING			
☐ SKEWED/SLANTED IMAGES			
☐ COLOR OR BLACK AND WHITE PHOTOGRAPHS			
☐ GRAY SCALE DOCUMENTS			
LINES OR MARKS ON ORIGINAL DOCUMENT			
☐ REFERENCE(S) OR EXHIBIT(S) SUBMITTED ARE POOR QUALITY			
OTHER:			

## IMAGES ARE BEST AVAILABLE COPY.

As rescanning these documents will not correct the image problems checked, please do not report these problems to the IFW Image Problem Mailbox.